

# Error Correction for Differential Linear Network Coding in Slowly-Varying Networks

Sven Puchinger<sup>◊</sup>, Michael Cyran\*, Robert F. H. Fischer<sup>◊</sup>, Martin Bossert<sup>◊</sup>, Johannes B. Huber\*

<sup>◊</sup>Institute of Communications Engineering, Ulm University, Germany

{sven.puchinger | robert.fischer | martin.bossert}@uni-ulm.de

\*Institute for Information Transmission, University of Erlangen-Nuremberg, Germany

{cyran | huber}@LNT.de

**Abstract**—Differential linear network coding (DLNC) is a precoding scheme for information transmission over random linear networks. By using differential encoding and decoding, the conventional approach of lifting, required for inherent channel sounding, can be omitted and in turn higher transmission rates are supported. However, the scheme is sensitive to variations in the network topology. In this paper, we derive an extended DLNC channel model which includes slow network changes. Based on this, we propose and analyze a suitable channel coding scheme matched to the situation at hand using rank-metric convolutional codes.

**Index Terms**—differential linear network coding, random linear network coding, rank-metric codes, partial-unit-memory codes

## I. INTRODUCTION

*Network coding*, introduced 2000 in [1], is a promising method for transmitting information over a network. [2] proved in 2003 that *linear network coding* (LNC) is a *max-flow* achieving approach for general multisource multicast networks. In LNC, the packets within the networks are vectors over a finite field  $\mathbb{F}_q$  and intermediate nodes transmit  $\mathbb{F}_q$ -linear combinations of the incoming packets. In 2006, [3] showed that choosing the linear functions at the nodes of a network in a random fashion achieves the *max-flow bound* with probability exponentially approaching 1 with the code length. This method is called *random linear network coding* (RLNC).

In this paper, we consider a unicast scenario, where one source node with  $n$  outgoing edges wants to transmit information to one destination node with  $\tilde{n} \geq n$  incoming edges over a network whose topology is neither known by the source nor by the destination. We assume that the transmission takes place in *generations*, which happen infinitely fast, or equivalently, every node of the network waits until all incoming edges have sent their packets before sending the outgoing packets. The source emits  $n$  vectors from  $\mathbb{F}_q^M$  (length- $M$  packets) into the network in parallel and the destination collects  $\tilde{n} \geq n$  such packets from the network during one generation. Within the network, nodes send—possibly different— $\mathbb{F}_q$ -linear combinations of the packets of their incoming edges to the outgoing edges. The vectors sent by the source during one generation can be seen as a matrix  $\mathbf{X} \in \mathbb{F}_q^{n \times M}$  and the collected packets by the receiver analogously as  $\mathbf{Y} \in \mathbb{F}_q^{\tilde{n} \times M}$ , where the rows of the matrices correspond to the packets. Errors are considered to be vectors added to nodes as additional incoming edges. This channel

This work was supported by Deutsche Forschungsgemeinschaft (DFG) within the framework COIN under grants BO 867/29-3, FI 982/4-3, and HU 634/11-3.

model is equivalent to the operator channel described in [4] under the assumption that the network stays constant during one generation. The input-output relation of the considered scenario is referred to as the *multiplicative additive matrix channel* (MAMC) [5]

$$\mathbf{Y} = \mathbf{A} \cdot \mathbf{X} + \mathbf{B}, \quad (1)$$

where  $\mathbf{A} \in \mathbb{F}_q^{\tilde{n} \times n}$  is called the *network channel matrix*, and  $\mathbf{B} \in \mathbb{F}_q^{\tilde{n} \times M}$  is the *additive error matrix*. It is commonly assumed (e.g., [5]) that  $\tilde{n} = n$  and  $\mathbf{A}$  is uniformly distributed at random among all regular matrices from  $\mathbb{F}_q^{n \times n}$ . The latter is a good assumption for sufficiently large networks and field sizes. We also assume that errors occur additively at arbitrary intermediate nodes. In this case, it can be shown (cf. [4]) that the rank of  $\mathbf{B}$  is upper bounded by the number of additive errors within the network during one generation.

Error correction for MAMCs can be done using *lifted rank-metric codes*. Alternatively, [6] showed that higher rates can be achieved by using differential precoding instead of lifting. This noncoherent transmission scheme for MAMCs is called *differential linear network coding* (DLNC). However, the method requires the channel matrix to remain constant between generations. This is a strong assumption which does not hold for all real-world communication networks. Hence, it has to be determined which influence a “slow” variation of the network (change of the multiplicative matrix in the MAMC) has on the error structure. We show that slow (i.e., the probability of a leaving/joining node during one generation is small) changes of the network topology result in impulsive error peaks. For this case, we propose a suitable channel coding scheme based on *rank-metric partial-unit-memory (PUM) codes* and analyze it in terms of transmission rates and error probabilities.

This paper is organized as follows: In Sec. II and III we briefly explain DLNC and rank-metric codes. An error model which includes slow changes of the network topology is introduced in Sec. IV and Sec. V shows a proper coding scheme for it and analyzes its performance. We conclude the paper with a short summary in Sec. VI.

We make use of the following notations. Let  $q$  be a prime power,  $m$  a positive integer. We write  $\mathbb{F}_q^{k \times \ell}$  for the set of all  $k \times \ell$  matrices over  $\mathbb{F}_q$ . Vectors are considered to be row vectors.  $\mathbf{I}_n$  denotes the identity matrix of size  $n \times n$ .  $\text{GL}(n, \mathbb{F}_q)$  is the set of all invertible  $(n \times n)$ -matrices over  $\mathbb{F}_q$ . The entry of a matrix  $\mathbf{A}$  in the  $i^{\text{th}}$  row and the  $j^{\text{th}}$  column is called  $A_{ij}$ . Let  $\beta := (\beta_1, \dots, \beta_m)$  be a basis<sup>1</sup> of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ . Then there

<sup>1</sup>E.g.,  $\beta = (1, \alpha, \alpha^2, \dots, \alpha^{m-1})$ , where  $\alpha$  is a primitive element of  $\mathbb{F}_{q^m}$ .

is a bijective linear map

$$\Phi_{\beta} : \mathbb{F}_{q^m}^n \rightarrow \mathbb{F}_q^{m \times n}, \mathbf{a} = (a_1, \dots, a_n) \mapsto \mathbf{A} = [A_{ij}],$$

where  $a_j = \sum_{i=1}^m A_{ij} \beta_i$  for all  $j \in \{1, \dots, n\}$ .  $f^{(*)\ell}(x)$  denotes the  $\ell$ -fold convolution of  $f(x)$ , i.e.,  $f^{(*)0}(x) = \delta(x)$ ,  $f^{(*)1}(x) = f(x)$ ,  $f^{(*)2}(x) = f(x) * f(x)$ ,  $f^{(*)3}(x) = f(x) * f(x) * f(x)$ , etc.

## II. DIFFERENTIAL LINEAR NETWORK CODING

In this section, we briefly describe the DLNC precoding method introduced in [6]. The concept of DLNC can be compared with *differential phase-shift keying*—information is not transmitted absolutely, but by the transition between two successive symbols.

We first restrict ourselves to square matrices, i.e.,  $n = M$ . For differential modulation, we assume to have a sequence of source words  $\mathbf{S}_i \in \text{GL}(n, \mathbb{F}_q)$ ,  $i = 1, \dots, N$ , for some  $N \in \mathbb{N}$ . Then we generate the DLNC transmit symbol  $\mathbf{X}_i$  of the  $i$ th generation as follows

$$\mathbf{X}_i = \mathbf{X}_{i-1} \cdot \mathbf{S}_i, \quad (2)$$

where  $\mathbf{X}_0 := \mathbf{I}_n$  is the initialization word.

Differential demodulation at the destination node starts by calculating the pseudoinverse  $\mathbf{Y}_{i-1}^+$  of the previously received matrix  $\mathbf{Y}_{i-1}$ , followed by calculating the product  $\mathbf{Y}_{i-1}^+ \cdot \mathbf{Y}_i$ . The pseudoinverse  $\mathbf{Y}_{i-1}^+$  has the following properties

$$\mathbf{Y}_{i-1}^+ \cdot \mathbf{Y}_{i-1} = \mathbf{I}_n + \mathbf{L} \mathbf{I}_{\mathcal{U}}^T, \quad (3)$$

$$\mathbf{Y}_{i-1} \cdot \mathbf{Y}_{i-1}^+ \cdot \mathbf{Y}_{i-1} = \mathbf{Y}_{i-1}. \quad (4)$$

Thereby,  $\mathbf{I}_{\mathcal{U}}$  consists of a subset  $\mathcal{U}$  of the rows of  $\mathbf{I}_n$ .  $\mathbf{L}$  is a full-rank matrix of appropriate dimensions and  $\mathbf{I}_{\mathcal{U}}^T \mathbf{L} = -\mathbf{I}_{|\mathcal{U}|}$ . Thus, with  $\mathbf{Y}_i = \mathbf{A}_i \mathbf{X}_i + \mathbf{B}_i$ , and if we assume that the network channel matrix  $\mathbf{A}_i$  stays constant between the two generations  $i-1$  and  $i$ , the demodulation process results in the demodulated symbol  $\tilde{\mathbf{S}}_i$

$$\begin{aligned} \tilde{\mathbf{S}}_i &= \mathbf{Y}_{i-1}^+ \mathbf{Y}_i = \mathbf{S}_i + \underbrace{\mathbf{L} \mathbf{I}_{\mathcal{U}}^T \mathbf{S}_i - \mathbf{Y}_{i-1}^+ \mathbf{B}_{i-1} \mathbf{S}_i + \mathbf{Y}_{i-1}^+ \mathbf{B}_i}_{=: \mathbf{E}_i} \\ &= \mathbf{S}_i + \mathbf{E}_i, \end{aligned} \quad (5)$$

where  $\mathbf{E}_i$  denotes the *effective error matrix* in generation  $i$ . It can be directly seen from (5) that differential modulation and demodulation transforms the MAMC with network channel matrix  $\mathbf{A}_i$  and additive error matrix  $\mathbf{B}_i$  into an *additive matrix channel* (AMC) [5] with additive error matrix  $\mathbf{E}_i$ . It is shown in [6] that  $\text{rk}(\mathbf{E}_i) \leq \text{rk}(\mathbf{B}_i) + \text{rk}(\mathbf{B}_{i-1})$ .

In case of non-square matrices  $\mathbf{X}_i$  ( $n \neq M$ ), the differential encoding reads  $\mathbf{X}_i = [\mathbf{X}_{i-1}]_{[n]} \cdot \mathbf{S}_i$ , where  $[\mathbf{X}_{i-1}]_{[n]} \in \text{GL}(n, \mathbb{F}_q)$  denotes the square matrix obtained from the first  $n$  columns of  $\mathbf{X}_{i-1}$ . The restriction to source matrices  $\mathbf{S}_i$  such that  $[\mathbf{X}_i]_{[n]}$  is invertible results in a rate loss of  $L_{\text{dlnc}} = \frac{n}{qM}$  (cf. [6]), which is negligible for sufficiently large field sizes  $q$ . Accordingly, differential demodulation is done via  $[\mathbf{Y}_{i-1}]^+_{[n]} \cdot \mathbf{Y}_i$ .

## III. RANK-METRIC PARTIAL-UNIT-MEMORY CODES

In this section, we give an overview of *rank-metric partial-unit-memory (PUM) codes*. We start by defining rank-metric block codes in general and Gabidulin codes in particular. Next, we review how rank-metric PUM codes are constructed in [7] and give an idea how to decode these codes and state a sufficient decoding condition (cf. [8], [9]).

### A. Rank-Metric Codes

Error-correcting codes are often assessed by the Hamming distance of their codewords. However, in 1978, Delsarte [10] introduced the so-called rank metric which turns out to be practical for certain channels, like the MAMC. The *rank weight* of an element  $\mathbf{a} \in \mathbb{F}_{q^m}^n$  is defined as  $\text{wt}_R(\mathbf{a}) := \text{rk}(\mathbf{A}) = \text{rk}(\Phi_{\beta}(\mathbf{a}))$ . Using this notation, we can define the *rank metric* of two elements  $\mathbf{a}, \mathbf{b} \in \mathbb{F}_{q^m}^n$  as  $d_R(\mathbf{a}, \mathbf{b}) := \text{wt}_R(\mathbf{a} - \mathbf{b}) = \text{rk}(\Phi_{\beta}(\mathbf{a}) - \Phi_{\beta}(\mathbf{b}))$ . The *minimum rank distance* of a code  $\mathcal{C}$  is defined as  $d := \min \{d_R(\mathbf{a} - \mathbf{b}) : \mathbf{a}, \mathbf{b} \in \mathcal{C} \wedge \mathbf{a} \neq \mathbf{b}\}$ .

As for the Hamming distance, a Singleton-like upper bound on the minimum rank distance can be given, i.e.,  $d \leq n - k + 1$  (cf. [10]–[12]). Codes fulfilling this bound with equality are called *maximum rank distance (MRD) codes*. A special class of MRD codes was introduced by Delsarte [10] and later independently reintroduced by Gabidulin [11] and Roth [12] and is usually called *Gabidulin codes*. Their structure and known decoding algorithms have a lot in common with *Reed–Solomon codes* in Hamming metric.

**Definition 1 (Gabidulin Code [11])** Let  $g_0, g_1, \dots, g_{n-1} \in \mathbb{F}_{q^m}$  be linearly independent over  $\mathbb{F}_q$ . Then a  $\mathcal{G}[n, k] \subset \mathbb{F}_{q^m}^n$  Gabidulin code is a linear code given by the following generator matrix:

$$\mathbf{G}_{\mathcal{G}} = \begin{bmatrix} g_0 & g_1 & \cdots & g_{n-1} \\ g_0^{q^1} & g_1^{q^1} & \cdots & g_{n-1}^{q^1} \\ \vdots & \vdots & \ddots & \vdots \\ g_0^{q^{k-1}} & g_1^{q^{k-1}} & \cdots & g_{n-1}^{q^{k-1}} \end{bmatrix} \quad (6)$$

A proof that these codes are MRD can be found in [11].

### B. Partial-Unit-Memory Codes in Rank Metric

We are considering convolutional codes in rank metric. In general, a convolutional code can be described by a semi-infinite block-Toeplitz generator matrix [13]. We only consider terminated generator matrices, which is not a restriction of generality in our case because we are dealing with finite information sequences. Such a generator matrix of a rate  $R = \frac{k}{n}$  convolutional code of memory  $\mu$  is given by

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}^{(0)} & \mathbf{G}^{(1)} & \cdots & \mathbf{G}^{(\mu)} & \mathbf{0} \\ \mathbf{G}^{(0)} & \mathbf{G}^{(1)} & \cdots & \mathbf{G}^{(\mu)} & \\ \ddots & \ddots & \ddots & \ddots & \\ \mathbf{0} & & & & \mathbf{G}^{(0)} & \mathbf{G}^{(1)} & \cdots & \mathbf{G}^{(\mu)} \end{bmatrix}, \quad (7)$$

where  $\mathbf{G}^{(i)} \in \mathbb{F}_q^{k \times n}$  for all  $i$  (cf. [13]).

PUM codes are convolutional codes of memory  $\mu = 1$ , introduced by Lee [14] and Lauer [15]. It can be shown (e.g. [16, Thm. 8.28]) that any convolutional code can be represented as a PUM code. However, PUM codes are usually constructed using block codes to obtain a good algebraic understanding of the convolutional code. For the following definition, let  $k^{(1)}, k \in \mathbb{N}$ , such that  $k^{(1)} \leq k \leq n - k^{(1)}$ .

**Definition 2** A  $\mathcal{PUM}(n, k, k^{(1)})$  code over  $\mathbb{F}_q$  is a rate  $\frac{k}{n}$  convolutional code with memory  $\mu = 1$  and generator matrix submatrices  $\mathbf{G}^{(0)}$  and  $\mathbf{G}^{(1)}$  having  $\text{rk}(\mathbf{G}^{(0)}) = k$  and  $\text{rk}(\mathbf{G}^{(1)}) = k^{(1)}$ .

W.l.o.g. we can assume that only the first  $k^{(1)}$  rows of  $\mathbf{G}^{(1)}$  are nonzero (otherwise we can transform the information sequence such that it has this form) and therefore we can subdivide the matrices as follows:

$$\mathbf{G}^{(0)} = \begin{bmatrix} \mathbf{G}^{(00)} \\ \mathbf{G}^{(01)} \end{bmatrix}, \quad \mathbf{G}^{(1)} = \begin{bmatrix} \mathbf{G}^{(10)} \\ \mathbf{0} \end{bmatrix}, \quad (8)$$

where  $\mathbf{G}^{(00)}, \mathbf{G}^{(10)} \in \mathbb{F}_q^{k^{(1)} \times n}$  and  $\mathbf{G}^{(01)} \in \mathbb{F}_q^{(k-k^{(1)}) \times n}$ .

It is shown in [7] and [8] how PUM codes can be constructed based on block rank-metric codes. One can construct PUM codes based on Gabidulin codes by choosing  $\mathbf{G}^{(00)}$ ,  $\mathbf{G}^{(10)}$  and  $\mathbf{G}^{(01)}$  as submatrices of generator matrices of Gabidulin code (7). In particular, the submatrices are chosen such that the codes in Tab. I are Gabidulin codes with the properties given in the table. We also choose a *bounded minimum distance (BMD)* error-erasure decoder for each of the defined codes, e.g., from [4] or [17].

Table I  
CODE DEFINITIONS USING SUB-MATRICES OF THE PUM CODE  
GENERATOR MATRIX [8]

Code	Generator Matrix	Type	Minimum Rank Distance	Decoder
$\mathcal{C}_0$	$\mathbf{G}^{(0)} = \begin{bmatrix} \mathbf{G}^{(00)} \\ \mathbf{G}^{(01)} \end{bmatrix}$	$\mathcal{G}[n, k]$	$d_0 = n - k + 1$	BMD( $\mathcal{C}_0$ )
$\mathcal{C}_1$	$\begin{bmatrix} \mathbf{G}^{(01)} \\ \mathbf{G}^{(10)} \end{bmatrix}$	$\mathcal{G}[n, k]$	$d_1 = n - k + 1$	BMD( $\mathcal{C}_1$ )
$\mathcal{C}_{01}$	$\mathbf{G}^{(01)}$	$\mathcal{G}[n, k - k^{(1)}]$	$d_{01} = n - k + k^{(1)} + 1$	BMD( $\mathcal{C}_{01}$ )
$\mathcal{C}_\sigma$	$\mathbf{G}^{(\sigma)} := \begin{bmatrix} \mathbf{G}^{(00)} \\ \mathbf{G}^{(01)} \\ \mathbf{G}^{(10)} \end{bmatrix}$	$\mathcal{G}[n, k + k^{(1)}]$	$d_\sigma = n - k - k^{(1)} + 1$	BMD( $\mathcal{C}_\sigma$ )

### C. BMD Decoding of Partial-Unit-Memory Codes

We consider codewords  $\mathbf{c}$  of PUM codes of length  $nN$ , obtained from an information sequence  $\mathbf{i}$  of length  $k(N-1)$  by multiplication with the (terminated) PUM generator matrix (7). We can divide the codeword  $\mathbf{c} \in \mathbb{F}_{q^m}^{nN}$  into  $N$  blocks of length  $n$ ,  $\mathbf{c} = (\mathbf{c}^{(1)} \dots \mathbf{c}^{(N)})$ , as well as the information word  $\mathbf{i} \in \mathbb{F}_{q^m}^{k(N-1)}$  into  $(N-1)$  blocks of length  $k$ ,  $\mathbf{i} = (\mathbf{i}^{(1)} \dots \mathbf{i}^{(N-1)})$ . With that notation, we can derive the simple encoding rule:

$$\mathbf{c}^{(j)} = \mathbf{G}^{(1)} \mathbf{i}^{(j-1)} + \mathbf{G}^{(0)} \mathbf{i}^{(j)} \quad \forall j \in \{1, \dots, N\}, \quad (9)$$

with  $\mathbf{i}^{(0)} = \mathbf{i}^{(N)} = \mathbf{0}$ . We use the decoder to correct an additive rank error  $\mathbf{e} = (\mathbf{e}^{(1)} \dots \mathbf{e}^{(N)}) \in \mathbb{F}_{q^m}^{nN}$ , so the received word is  $\mathbf{r} = \mathbf{c} + \mathbf{e} = (\mathbf{r}^{(1)} \dots \mathbf{r}^{(N)}) \in \mathbb{F}_{q^m}^{nN}$ .

It is known that convolutional codes can be ML decoded using the *Viterbi* algorithm [18]. However, the complexity of this procedure depends strongly on the size of the underlying field, which defines the number of states of the respective trellis. The necessary field size over which a Gabidulin code must be defined grows exponentially with the codelength  $n$ . Therefore, Viterbi's algorithm is not a good choice for decoding PUM codes based on Gabidulin codes.

An alternative is an algorithm introduced by Dettmar and Sorger [19] for decoding PUM codes over Hamming metric. It uses the block decoders of the underlying algebraic codes to find a much smaller subgraph of the trellis which contains the most likely code sequence under a certain condition.

Afterwards, the Viterbi algorithm finds this sequence in the reduced trellis in less time than without the reduction step.

In [8], a generalization of this algorithm to rank metric was proposed. It is shown that the following bound provides a sufficient condition for successful decoding, using the rank weight distribution of the additive error  $\mathbf{e}$ .

$$\sum_{h=i}^{i+j-1} t^{(h)} < \frac{\delta_j}{2}, \quad \begin{cases} \forall i \in \{0, \dots, N\} \text{ and} \\ \forall j \in \{1, \dots, N-i+1\}, \end{cases} \quad (10)$$

where  $t^{(h)} := \text{wt}_R(\mathbf{E}^{(h)}) = \text{wt}_R(\Phi_\beta(\mathbf{e}^{(h)}))$  and

$$\delta_j := \begin{cases} d_{01}, & \text{if } j = 1, \\ d_0 + (j-2)d_\sigma + d_1, & \text{else.} \end{cases} \quad (11)$$

If the sequence is sent block-by-block using a Gabidulin code with dimension  $k$  instead, all error patterns containing at least one block  $h$  with  $t^{(h)} \geq \frac{d_0}{2}$  will not be decoded correctly. PUM codes can handle such *error peaks* if  $t^{(h)} < \frac{d_{01}}{2}$  and if the errors in the surrounding blocks are not too large (cf. (10)). This property makes them suitable for DLNC in slowly-varying networks because they can handle seldom occurring network changes resulting in impulsive error peaks. It can be shown that in terms of the *sequence error probability*  $\text{Pr}_{\text{fail}}$ , i.e., the probability that the decoder fails for at least one block of the sequence, the PUM decoder is strictly better than the block-by-block decoder.

### IV. DLNC CHANNEL MODEL FOR SLOWLY-VARYING NETWORKS

We analyze the statistical behavior of the rank of the effective error matrix  $\mathbf{E}$ , cf., (5). For that purpose, we need to discuss the statistical behavior of the additive error matrix  $\mathbf{B}$ , and the effects of a slowly-varying network topology. In order to derive an analytic expression for the probability distributions, we make the following assumptions:

- 1) The probability  $p_n$  that an error occurs at a certain node during one generation is constant.
- 2) The probability  $p_{\Delta n}$  that a certain node leaves or joins the network during one generation is constant.
- 3) The probability  $p_e$  that there is a directed edge from node  $i$  to  $j$  ( $i \neq j$ ) is constant.
- 4) The number of nodes  $|\mathcal{N}|$  is sufficiently large and can be assumed to be constant due to very slow network changes.
- 5) The field size  $q$  is sufficiently large such that the probability that two independent errors cancel is close to zero.

#### A. Additive Error Matrix

It was shown in [6] that the rank of the additive error matrix  $\mathbf{B}$  is approximately distributed binomially with parameters  $|\mathcal{N}|$  and  $p_n$  if Assumption 5 holds. Hence, its *probability mass function* (pmf) with mean value  $\mu_B = |\mathcal{N}| \cdot p_n$  is given by

$$f_{\text{rk}(\mathbf{B})}(\tau) \approx \binom{|\mathcal{N}|}{\tau} \cdot p_n^\tau \cdot (1 - p_n)^{|\mathcal{N}| - \tau}. \quad (12)$$

#### B. Slow Changes of the Network Topology

If nodes leave or join the network between generations, the network behavior changes and can be expressed as a difference in the channel matrices

$$\mathbf{A}_i = \mathbf{A}_{i-1} + \Delta \mathbf{A}_i, \quad (13)$$

where we call  $\Delta\mathbf{A}_i$  the *channel deviation*. Since the rank of  $\Delta\mathbf{A}_i$  turns out to be important for the effective error matrix using DLNC, we are interested in its distribution. We start by proving an important theorem.

**Theorem 1** *If exactly one node  $\nu$  with  $\text{In}\{\nu\}$  incoming and  $\text{Out}\{\nu\}$  outgoing edges leaves the network between generation  $i-1$  and  $i$ , the rank of  $\Delta\mathbf{A}_i$  is upper bounded by  $\text{rk}(\Delta\mathbf{A}_i) \leq \min\{n, \text{Out}\{\nu\}, \text{In}\{\nu\}\}$ .*

**Proof** Due to the dimensions of  $\Delta\mathbf{A}_i$ , its rank is upper bounded by  $n$ . Let the outgoing packets be independent random linear combinations  $\mathbf{x}_{o,j}$  of the incoming edges, sent to nodes  $\nu_{o,j}$  with  $j \in \{1, \dots, \text{Out}\{\nu\}\}$ . The channel deviation can then be interpreted as errors of value  $e_{o,j} := -\mathbf{x}_{o,j}$  at each node  $\nu_{o,j}$ . By the same argument as for the additive error matrix (cf. [4]), the rank of  $\Delta\mathbf{A}_i$  is then upper bounded by the number of these additive errors, namely  $\text{Out}\{\nu\}$ . Alternatively, the channel deviation can be seen as additive errors at the origin nodes of the  $\text{In}\{\nu\}$  incoming edges. Hence,  $\text{rk}(\Delta\mathbf{A}_i) \leq \text{In}\{\nu\}$ . The proof is illustrated in Fig. 1. ■

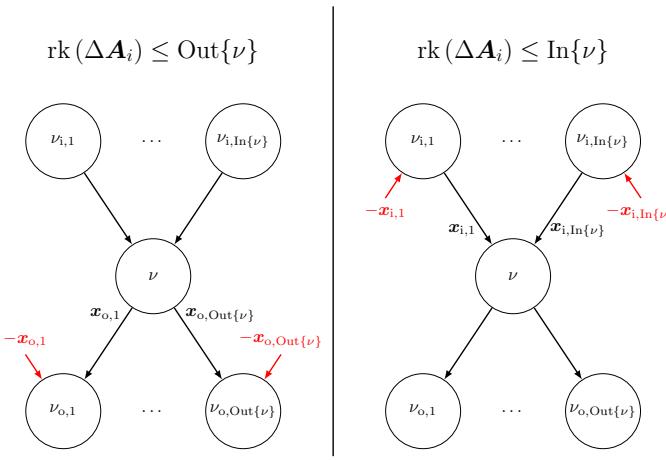


Figure 1. Illustration of the proof of Theorem 1.

The same argument holds for nodes that join the network between generations. For a leaving or joining node  $\nu$ , we define  $w(\nu) := \min\{n, \text{Out}\{\nu\}, \text{In}\{\nu\}\}$ , and  $\ell_i$  to be the number of leaving nodes in generation  $i$ . Its pmf  $f_W(w)$  can be derived using Assumption 3. Due to the subadditivity of the rank, we obtain the following upper bound on  $\text{rk}(\Delta\mathbf{A}_i)$ .

$$\text{rk}(\Delta\mathbf{A}_i) \leq \sum_{\substack{\nu \text{ leaving} \\ \text{in gen. } i}} w(\nu) = \sum_{j=1}^{\ell_i} w(\nu_j)$$

This means that the rank of the deviation matrix  $\Delta\mathbf{A}_i$  is composed by two random processes. The first one determines the number  $\ell_i$  of leaving/joining nodes before the current generation according to the pmf  $f_L(\ell)$ , which is binomially distributed with  $|\mathcal{N}|$  and  $p_{\Delta n}$ . The second process determines the  $\ell_i$  corresponding node weights  $w(\nu_j)$ , which are distributed according to  $f_W(w)$ . Hence, the pmf  $f_{\text{rk}(\Delta\mathbf{A}_i)}$  is approximately given by

$$f_{\text{rk}(\Delta\mathbf{A}_i)}(\tau) = \sum_{\ell=0}^{|\mathcal{N}|} f_L(\ell) \cdot f_W^{(\ast)\ell}(\tau). \quad (14)$$

### C. Effective Error Matrix

Based on the insights of the last two paragraphs, we are able to bound the rank of the effective error matrix  $\mathbf{E}_i$ .

**Theorem 2** *The rank of the effective error matrix  $\mathbf{E}_i$ , which is present in a DLNC system applied to a slowly-varying network is upper bounded by*

$$\text{rk}(\mathbf{E}_i) \leq \text{rk}(\mathbf{B}_{i-1}) + \text{rk}(\mathbf{B}_i) + \text{rk}(\Delta\mathbf{A}). \quad (15)$$

Given Assumption 5, the bound is tight with high probability.

**Proof** Inserting (1) and (13) into (5) results in

$$\begin{aligned} \tilde{\mathbf{S}}_i &= \mathbf{Y}_{i-1}^+ [(\mathbf{A}_{i-1} + \Delta\mathbf{A}_i)\mathbf{X}_i + \mathbf{B}_i] \\ &= \mathbf{Y}_{i-1}^+ \mathbf{A}_{i-1} \mathbf{X}_i + \mathbf{Y}_{i-1}^+ \Delta\mathbf{A}_i \mathbf{X}_i + \mathbf{Y}_{i-1}^+ \mathbf{B}_i. \end{aligned} \quad (16)$$

With the aid of (3)  $\mathbf{Y}_{i-1} = \mathbf{I}_n + \mathbf{L}\mathbf{I}_{\mathcal{U}}^T$ , we can rearrange  $\mathbf{Y}_{i-1} = \mathbf{A}_{i-1} \mathbf{X}_{i-1} + \mathbf{B}_{i-1}$  in the following way

$$\mathbf{Y}_{i-1}^+ \mathbf{A}_{i-1} = \mathbf{X}_{i-1}^{-1} + \mathbf{L}\mathbf{I}_{\mathcal{U}}^T \mathbf{X}_{i-1}^{-1} - \mathbf{Y}_{i-1}^+ \mathbf{B}_{i-1} \mathbf{X}_{i-1}^{-1}. \quad (17)$$

Combining (16) and (17) we obtain

$$\begin{aligned} \tilde{\mathbf{S}}_i &= \left( \mathbf{X}_{i-1}^{-1} + \mathbf{L}\mathbf{I}_{\mathcal{U}}^T \mathbf{X}_{i-1}^{-1} - \mathbf{Y}_{i-1}^+ \mathbf{B}_{i-1} \mathbf{X}_{i-1}^{-1} \right) \mathbf{X}_i \\ &\quad + \mathbf{Y}_{i-1}^+ \Delta\mathbf{A}_i \mathbf{X}_i + \mathbf{Y}_{i-1}^+ \mathbf{B}_i \\ &= \mathbf{X}_{i-1}^{-1} \mathbf{X}_i + \mathbf{L}\mathbf{I}_n^T \mathbf{X}_{i-1}^{-1} \mathbf{X}_i - \mathbf{Y}_{i-1}^+ \mathbf{B}_{i-1} \mathbf{X}_{i-1}^{-1} \mathbf{X}_i \\ &\quad + \mathbf{Y}_{i-1}^+ \Delta\mathbf{A}_i \mathbf{X}_i + \mathbf{Y}_{i-1}^+ \mathbf{B}_i \\ &= \mathbf{S}_i + \underbrace{\mathbf{L}\mathbf{I}_{\mathcal{U}}^T \mathbf{S}_i - \mathbf{Y}_{i-1}^+ \mathbf{B}_{i-1} \mathbf{S}_i + \mathbf{Y}_{i-1}^+ \mathbf{B}_i + \mathbf{Y}_{i-1}^+ \Delta\mathbf{A}_i \mathbf{X}_i}_{=: \mathbf{E}_i}. \end{aligned}$$

Thus, the effective error matrix consists of four parts. The first part corresponds to the rank deficiency of  $\mathbf{Y}_{i-1}$ , which is assumed to be zero. This assumption is justifiable as long as  $q$  is sufficiently large (cf. [6]). The second and the third part have the same rank as the preceding and the current additive error matrices,  $\mathbf{B}_{i-1}$  and  $\mathbf{B}_i$ , respectively. The rank of the last part equals to  $\text{rk}(\Delta\mathbf{A}_i)$ . Finally, taking the subadditivity of the rank into account, we obtain (15) as an upper bound on the rank of  $\mathbf{E}_i$ . Due to Assumption 5, with high probability, the sum of the matrices has the same rank as the sum of the ranks. ■

As a consequence, the rank  $\mathbf{E}_i$  can approximately be described by the sum of three random variables

$$\text{rk}(\mathbf{E}_i) = \text{rk}(\mathbf{B}_{i-1}) + \text{rk}(\mathbf{B}_i) + \text{rk}(\Delta\mathbf{A}), \quad (18)$$

where the first two summands, which describe the effect of the additive error in the MAMC are distributed according to (12). The third summand describes the effect of the slowly-varying network and its pmf is given by (14). The resulting approximate pmf can be described as

$$f_{\text{rk}(\mathbf{E}_i)}(\tau) = f_{\text{rk}(\mathbf{B})}^{(\ast)2}(\tau) * f_{\text{rk}(\Delta\mathbf{A})}(\tau). \quad (19)$$

**Example 1** Fig. 2 illustrates the influence of the additive rank error and the rank error caused by slow network changes. The plot shows the pmfs of the ranks of the additive error  $\mathbf{B}_i$ , the channel deviation  $\Delta\mathbf{A}_i$ , the additive error without the influence of network changes  $\text{rk}(\tilde{\mathbf{E}}_i) := \text{rk}(\mathbf{B}_i) + \text{rk}(\mathbf{B}_{i-1})$  and the effective error  $\mathbf{E}_i$ . We consider a network with  $|\mathcal{N}| = 100$  nodes and probability parameters  $p_n = 0.03$ ,  $p_e = 0.05$  and  $p_{\Delta n} = 0.01$ . We have chosen  $p_{\Delta n}$  relatively large such

that the changes are more visible in the illustration. Plot (iii) in Fig. 2 depicts the effect caused by the differential demodulation in case of no network changes (cf. [6]), i.e., the expected value of the rank of the effective error is approximately doubled with respect to the additive error given by the MAMC (plot (i)). The effect of additional slow network changes can be seen in plot (iv). In contrast to (iii), both variance and mean are increased, making errors with high rank more likely. These error peaks can be better handled using PUM codes than using block codes.

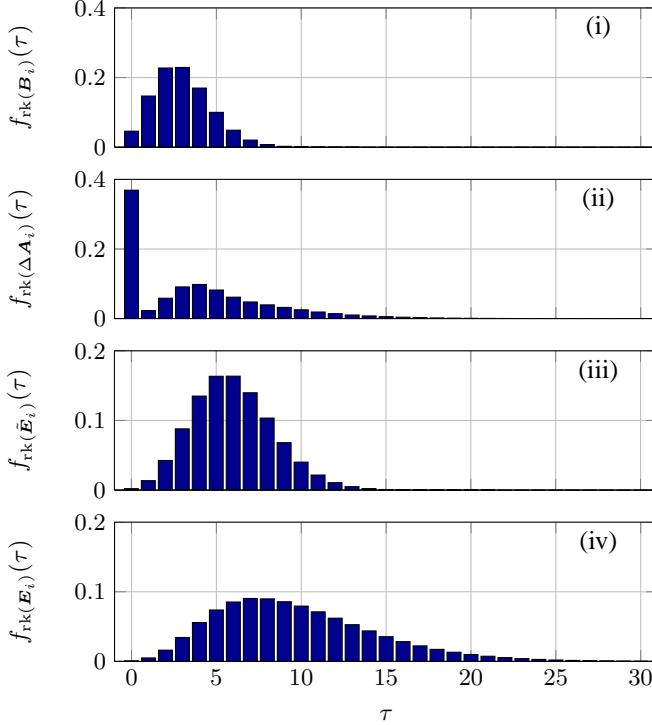


Figure 2. Influence of slow network changes on the pmf of  $\text{rk}(\mathbf{E}_i)$ .

## V. SUITABLE CODING SCHEMES

In this section, we describe how PUM codes can be used to make DLNC resilient against the additive error peaks caused by network changes. We show how PUM codes can be combined with DLNC, give an idea how to choose code parameters and present numerical results which demonstrate that better results can be obtained using PUM codes instead of ordinary block codes.

### A. Combining DLNC and PUM Codes

An information sequence  $\mathbf{i}^{(j)} \in \mathbb{F}_{q^m}^k$  for  $j = 1, \dots, N-1$  has to be transmitted using PUM codes in combination with DLNC. We first determine the corresponding codeword sequence  $\mathbf{c}^{(j)} \in \mathbb{F}_{q^m}^n$  for  $j = 1, \dots, N$  using the encoding rule (9) and calculate the corresponding matrix representation sequence  $\mathbf{S}_j := \Phi_{\beta}(\mathbf{c}^{(j)})$  for  $j = 1, \dots, N$ , which we use as source symbols for the differential encoding described in Sec. II. After the transmission, the sequence is demodulated and the result  $\tilde{\mathbf{S}}_j = \mathbf{S}_j + \mathbf{E}_j$  ( $j = 1, \dots, N$ ) can be decoded using the PUM BMD decoder described in Sec. III-C. The error sequence  $\mathbf{E}_j$  is distributed according to the error model derived in Sec. IV.

### B. PUM Code Parameter Choices

When designing a PUM code for DLNC in slowly-varying networks, one has several possibilities to choose the code parameters  $N$  and  $k^{(1)}$ . Here, we assume that  $n$  and  $k$  are fixed, e.g., because the desired code rate and the packet size is given. In general,  $N$  should be chosen sufficiently large, such that the rate loss of the termination of the PUM code does not play a role, i.e.,  $\frac{N}{N+1} \cdot \frac{k}{n} \approx \frac{k}{n}$ .

It is not easy to analytically derive a good range for  $k^{(1)}$  for general DLNC channels in slowly-varying networks due to the involved analytic description of the pmf of the effective rank error (19). However, for a given network, one can use the pmf of  $\text{rk}(\mathbf{E}_i)$  to get an idea how to choose  $k^{(1)}$ . As already mentioned, the PUM decoder described in [8] is able to decode up to  $\lfloor \frac{d_{01}-1}{2} \rfloor$  errors under certain conditions. Hence, a necessary condition for good decoding results is that  $\Pr \{ \text{rk}(\mathbf{E}_i) \geq \frac{d_{01}}{2} = \frac{n-k+k^{(1)}+1}{2} \}$  is small, or equivalently, that  $k^{(1)}$  is as large as possible. On the other hand,  $k^{(1)}$  should not be chosen too large, because otherwise  $d_{\sigma} = n-k-k^{(1)}+1$  gets too small and the decoding capabilities decrease again (cf. (10)). The following figure illustrates this behavior by showing how  $\text{Pr}_{\text{fail}}$  changes as a function of  $k^{(1)}$  for given  $p_n = 0.03$ ,  $p_e = 0.05$ ,  $p_{\Delta n} = 0.005$ ,  $|\mathcal{N}| = 100$ ,  $N = 50$  and  $d_0 = n-k+1 = 41$ . It can be seen from Fig. 3 that  $k^{(1)} = 24$  is an optimal choice for this given parameter set.

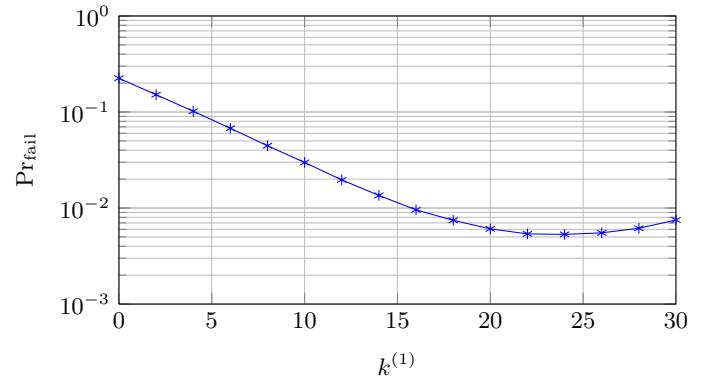


Figure 3.  $\text{Pr}_{\text{fail}}$  as a function of  $k^{(1)}$ .

### C. Numerical Results

Example 2 shows numerical results obtained by simulating random linear networks with given parameters  $|\mathcal{N}|$ ,  $p_n$  and  $p_e$  (cf. Sec. IV). In order to run the simulations in sufficiently short time<sup>2</sup>, we assume that the field size  $q$  is large enough, such that independent errors cancel only with negligibly small probability, and therefore the ranks of the error matrices are very likely to be equal to the number of errors happened. If this assumption does not hold (e.g., if  $q$  is relatively small), our results are still upper bounds on the sequence error probability. We used rank-metric PUM codes with parameters  $N$ ,  $n$ ,  $k$  and  $k^{(1)}$  and checked if the PUM decoder, described in Sec. III-C, is able to correct the error pattern.  $\text{Pr}_{\text{fail}}$  denotes the probability that the PUM decoder fails, i.e., at least one generation of the

<sup>2</sup>Instead of simulating RLNC using real networks over finite fields, we evaluated the statistical behavior of the nodes and counted the number of additive errors and leaving/joining nodes and their number of incoming and outgoing edges.

sequence is not contained in the subgraph of the trellis. For comparison, we also checked the cases (block-by-block) when the source symbols for every generation are encoded using Gabidulin block codes with the same code dimension  $k$ , both differentially and via lifting.

**Example 2** Fig. 4 shows the sequence failure probability  $\Pr_{\text{fail}}$  as a function of the network change probability  $p_{\Delta n}$  of a PUM decoder of a code with  $N = 50$ ,  $n = 35$ ,  $k = 15$  and  $k^{(1)} = 10$  and compares it to block-by-block decoders in combination with both DLNC and lifting. The network is assumed to have  $|\mathcal{N}| = 100$  nodes and probability parameters  $p_n = 0.01$  and  $p_e = 0.02$ . It can be seen that the PUM decoder is better than the block-by-block decoder for any  $p_{\Delta n}$  and also improves upon the lifting approach for  $p_{\Delta n} \leq 10^{-2}$ .

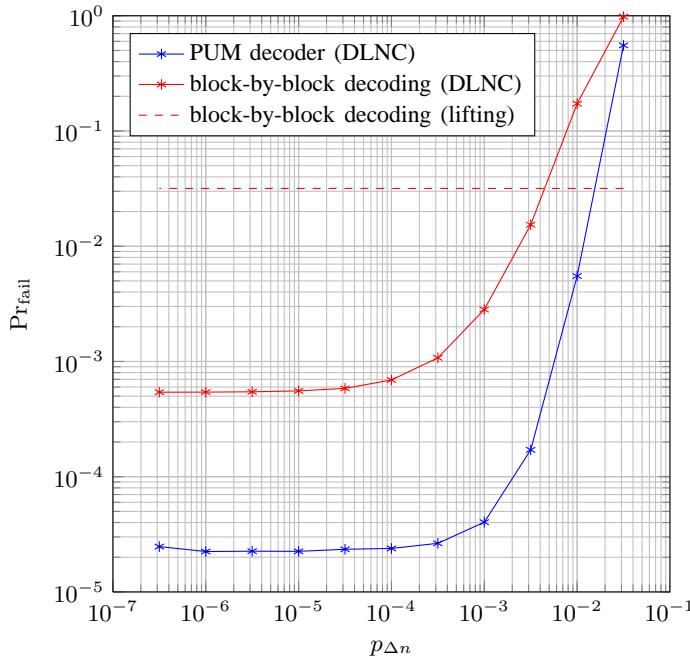


Figure 4.  $\Pr_{\text{fail}}$  of different codes/decoders as a function of  $p_{\Delta n}$ .

Fig. 5 shows the gain obtained by the new coding scheme, i.e., the fraction of  $\Pr_{\text{fail}}$  of the block-by-block decoder for DLNC and the PUM decoder. It makes clear that not only the PUM decoder is better at any  $p_{\Delta n}$ , but especially good in a region where  $p_{\Delta n}$  is relatively large ( $3 \cdot 10^{-4}$  to  $10^{-2}$ ).

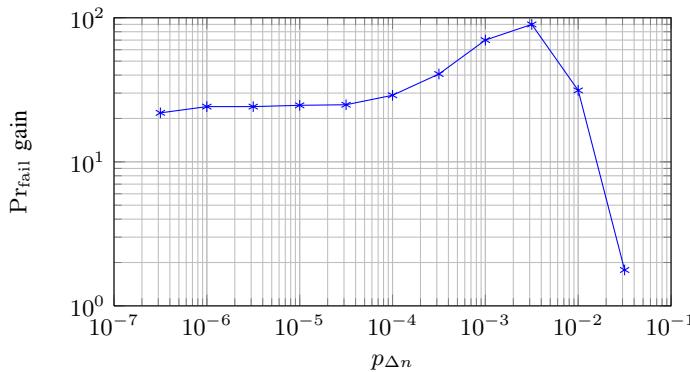


Figure 5. Gain of  $\Pr_{\text{fail}}$  of PUM codes compared to block-by-block decoding.

## VI. CONCLUSION

This paper extended the results of [6] to the case of slowly-varying networks. At first, we derived a probabilistic DLNC channel model for this case by analyzing the effects of joining/leaving nodes on the network channel matrix, and with that, on the effective error (5) in a DLNC system. Furthermore, we showed that PUM rank-metric codes are the proper error correction strategy for the situation at hand. We confirmed our considerations by numerical simulations, and showed that in slowly-varying networks, DLNC in combination with rank-metric PUM codes outperforms the conventional, lifting-based RLNC approach.

## REFERENCES

- [1] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.
- [2] S.-Y. Li, R. W. Yeung, and N. Cai, "Linear Network Coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003.
- [3] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A Random Linear Network Coding Approach to Multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.
- [4] R. Koetter and F. R. Kschischang, "Coding for Errors and Erasures in Random Network Coding," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3579–3591, 2008.
- [5] D. Silva, F. R. Kschischang, and R. Köetter, "Communication Over Finite-Field Matrix Channels," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1296–1305, 2010.
- [6] M. Seidl, M. Cyran, R. F. H. Fischer, and J. B. Huber, "A Differential Encoding Approach to Random Linear Network Coding," in *Proc. 9. International ITG Conference on Systems, Communications and Coding (SCC)*, Munich, Germany, Jan. 2013.
- [7] A. Wachter, V. R. Sidorenko, M. Bossert, and V. V. Zyablov, "On (Partial) Unit Memory Codes Based on Gabidulin Codes," *Problems of Information Transmission*, vol. 47, no. 2, pp. 117–129, 2011.
- [8] A. Wachter-Zeh, M. Stinner, and V. Sidorenko, "Convolutional Codes in Rank Metric with Application to Random Network Coding," *arXiv preprint arXiv:1404.7251*, 2014.
- [9] A. Wachter-Zeh and V. Sidorenko, "Rank Metric Convolutional Codes for Random Linear Network Coding," in *IEEE International Symposium on Network Coding (NetCod)*, 2012.
- [10] P. Delsarte, "Bilinear Forms over a Finite Field, with Applications to Coding Theory," *Journal of Combinatorial Theory, Series A*, vol. 25, no. 3, pp. 226–241, 1978.
- [11] E. M. Gabidulin, "Theory of Codes with Maximum Rank Distance," *Problemy Peredachi Informatsii*, vol. 21, no. 1, pp. 3–16, 1985.
- [12] R. M. Roth, "Maximum-Rank Array Codes and Their Application to Crisscross Error Correction," *IEEE Transactions on Information Theory*, vol. 37, no. 2, pp. 328–336, 1991.
- [13] R. Johannesson and K. S. Zigangirov, *Fundamentals of Convolutional Coding*. Wiley-IEEE press, 1999.
- [14] L.-N. Lee, "Short Unit-Memory Byte-Oriented Binary Convolutional Codes Having Maximal Free Distance," *IEEE Transactions on Information Theory*, vol. 22, no. 3, pp. 349–352, 1976.
- [15] G. Lauer, "Some Optimal Partial-Unit-Memory Codes," *IEEE Transactions on Information Theory*, vol. 25, no. 2, pp. 240–243, 1979.
- [16] M. Bossert, *Channel Coding for Telecommunications*. New York, NY, USA: John Wiley & Sons, Inc., 1999.
- [17] E. M. Gabidulin and N. I. Pilipchuk, "Error and erasure correcting algorithms for rank codes," *Designs, codes and Cryptography*, vol. 49, no. 1-3, pp. 105–122, 2008.
- [18] A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," *IEEE Transactions on Information Theory*, vol. 13, no. 2, pp. 260–269, 1967.
- [19] U. Dettmar and U. K. Serger, "Bounded Minimum Distance Decoding of Unit Memory Codes," *IEEE Transactions on Information Theory*, vol. 41, no. 2, pp. 591–596, 1995.